

(1) GENERAL INFORMATION

(ii) TITLE OF THE INVENTION: SH3-CONTAINING PROTEINS

(iii) NUMBER OF SEQUENCES: 5

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Incyte Pharmaceuticals, Inc.

(B) STREET: 3174 Porter Dr.

(C) CITY: Palo Alto

(D) STATE: CA

(E) COUNTRY: USA

(F) ZIP: 94304

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Diskette

(B) COMPUTER: IBM Compatible

(C) OPERATING SYSTEM: DOS

(D) SOFTWARE: FastSEQ for Windows Version 2.0

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: To Be Assigned

(B) FILING DATE: Filed Herewith

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER:

(B) FILING DATE:

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Billings, Lucy J.

(B) REGISTRATION NUMBER: 36,749

(C) REFERENCE/DOCKET NUMBER: PF-0419 US

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 650-855-0555

(B) TELEFAX: 650-845-4166

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 265 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

(A) LIBRARY: BRAITUT03

(B) CLONE: 865744

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Met	Lys	Arg	Thr	Val	Ser	Asp	Asn	Ser	Leu	Ser	Asn	Ser	Arg	Gly	Glu
				5					10					15	
Gly	Lys	Pro	Asp	Leu	Lys	Phe	Gly	Gly	Lys	Ser	Lys	Gly	Lys	Leu	Trp
			20					25					30		
Pro	Phe	Ile	Lys	Lys	Asn	Lys	Gly	Ala	Thr	Pro	Glu	Asp	Phe	Ser	Asn
		35					40					45			
Leu	Pro	Pro	Glu	Gln	Arg	Arg	Lys	Lys	Leu	Gln	Gln	Lys	Val	Asp	Glu

(2) INFORMATION FOR SEQ ID NO:2:

(A) LENGTH: 1459 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(A) LIBRARY: BRAITUT03
(B) CLONE: 865744

AGTAAAGCA	GCCGAATCAA	TTGATCAGAA	AAATGATTCA	CAGCTGGTAA	TAGAAGCTTA	60
TAAATCAGGG	TTTGAGCCTC	CTGGAGACAT	TGAATTTGAG	GATTACACTC	AGCCAATGAA	120
GCGCACTGTG	TCAGATAACA	GCCTTTCAAA	TTCCAGAGGA	GAAGGCAAAC	CAGACCTCAA	180
ATTTGGTGGC	AAATCCAAAG	GAAAGTTATG	GCCGTTCATC	AAAAAAAAATA	AGGGTGCAAC	240
ACCGGAGGAT	TTCAGCAACC	TCCCACCTGA	ACAAAGGAAG	AAAAAGCTGC	AGCAGAAAGT	300
CGATGAGTTA	AATAAAGAAA	TTCAGAAGGA	GATGGATCAA	AGAGATGCCA	TAAACAAAAAT	360
GAAAGATGTC	TACCTAAAGA	ATCCTCAGAT	GGGAGACCCA	GCCAGTTTGG	ATCACAAATTT	420
AGCAGAAGTC	AGCCAAAATA	TAGAGAAACT	GCGAGTAGAG	ACCCAGAAAT	TTGAGGCCCTG	480
GCTGGCTGAG	GTTGAAGGCC	GGCTCCCAGC	ACGCAACGAG	CAGGCGCGCC	GGCAGAGCGG	540
ACTGTACGAC	AGCCAGAACC	CACCCACAGT	CAACAACCTG	GCCCAGGACC	GTGAGAGCCC	600
AGATGGCAGT	TACACAGAGG	AGCAGAGTCA	GCGAGGTGAG	ATGAAGGTGC	TGGCCACGGA	660
TTTTGACGAC	GAGTTTGATG	ATGAGGAGCC	CCTCCCTGCC	ATAGGGACGT	GCAAAGCTCT	720
CTACACATTT	GAAGGTCAGA	ATGAAGGAAC	GATTTCCGTA	GTTGAAGGAG	AAACATTTGTA	780
TGTCATAGAG	GAAGACAAAG	GCGATGGCTG	GACCCGCATT	CGGAGAAATG	AAGATGAAGA	840
GGGTATATGC	CCCACCTTCAT	ATGTCGAAGT	CTGTTTGGAC	AAAAATGCCA	AAGGTGCTAA	900
GACTTATATT	TAATACCATA	AAAAAAAAAA	ACTTAAAAAA	AATGGAGTTG	TTTCTCCCCA	960
CAACCGTAGC	GTTTACAGGC	AGTTTCTCAA	GAGACTGGCT	GGCAAGCACC	ATAATGCACG	1020
TTCTCCTGTA	GTCTCACGTG	GACTTCAGGG	TCCGGGCACC	TGAAATTGCC	TGCTTAGTTT	1080
GGGCTGTAAT	CAAGTTTCAC	TTGCTGATGA	AATTTTATGT	GGAAAGCTGC	CAACCGCCAA	1140
CTTACAGCTA	TGTCATTCAA	AATCTGATAA	ACATTTCTTC	TTTTGGCGGT	ATCTGTAGAT	1200
TAAAAAATA	GTTGCATTGT	AGCTTCTCAT	CTTTCTGAAT	TTAAAAGCCG	GCACGCATCA	1260
TGCAGGTGCC	AAAGACTTTG	CTACTCTTGT	TTATATCTAG	TATCCACCAT	ACACTGAGCT	1320
ACATTAGGTG	GTTACAGATT	GTAACCTTAAT	AAACTGAACT	GTGTTAGTTT	GTTAAATTGG	1380
ATACTCATTC	ACTTGGGGAG	GAGTCACAAG	TGAAATACCA	TCTCTTTCTT	GACTAAAGCG	1440

GTAAATAAGG TTCTTATTG

1459

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 175 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: PROSNOT20
 (B) CLONE: 1816529

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Met Lys Asp Val Tyr Glu Lys Thr Pro Gln Met Gly Asp Pro Ala Ser
 1 5 10 15
 Leu Glu Pro Gln Ile Ala Glu Thr Leu Ser Asn Ile Glu Arg Leu Lys
 20 25 30
 Leu Glu Val Gln Lys Tyr Glu Ala Trp Leu Ala Glu Ala Glu Ser Arg
 35 40 45
 Val Leu Ser Asn Arg Gly Asp Ser Leu Ser Arg His Ala Arg Pro Pro
 50 55 60
 Xaa Pro Pro Ala Ser Ala Pro Pro Asp Ser Ser Asn Ser Ala Ser
 65 70 75 80
 Gln Asp Thr Lys Glu Ser Ser Glu Glu Pro Pro Ser Glu Glu Ser Gln
 85 90 95
 Asp Thr Pro Ile Tyr Thr Glu Phe Asp Glu Asp Phe Glu Glu Glu Pro
 100 105 110
 Thr Ser Pro Ile Gly His Cys Val Ala Ile Tyr His Phe Glu Gly Ser
 115 120 125
 Ser Glu Gly Thr Ile Ser Met Ala Glu Gly Glu Asp Leu Ser Leu Met
 130 135 140
 Glu Glu Asp Lys Gly Asp Gly Trp Thr Arg Val Arg Arg Lys Glu Gly
 145 150 155 160
 Gly Glu Gly Tyr Val Pro Thr Ser Tyr Leu Arg Val Thr Leu Asn
 165 170 175

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 773 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: PROSNOT20
 (B) CLONE: 1816529

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

ATGAACCGTG CACCCTNCGA CAGCAGTCTG GGCACCCCTT ACGGATGGAC GGNCTGAACT 60
 CCGAGGNCCG GGTCGCAGCC GCACCAAGCG CTGGNCTTTT GGCAAGAAGA ACAAGACAGT 120
 GGTGACCGAG GATTTTAGCC ACTTGCCCCC AGAGCAGCAG CGAAAACGGC TTCAACAGCA 180
 GTTGAAGAA CGCAGTCGTG AACTTCAGAA GGAGGTTGAC CAGAGGGAAG CCCTAAAGAA 240
 AATGAAGGAT GTCTATGAGA AGACACCTCA GATGGGGGAC CCCGCCAGCT TGGAGCCCCA 300
 GATCGCTGAA ACCCTGAGCA ACATTGAACG GCTGAAATTG GAAGTGCAGA AGTATGAGGC 360
 GTGGCTGGCA GAAGCTGAAA GTCGAGTCCT TAGCAACCGG GGAGACAGCC TGAGCCGGCA 420
 CGCCCGGCCT CCCGANCCCC CCGCTAGCGC CCCGCCAGAC AGCAGCAGCA ACAGCGCATC 480
 ACAGGACACC AAGGAGAGCT CTGAAGAGCC TCCCTCAGAA GAGAGCCAGG ACACCCCCAT 540
 TTACACGGAG TTTGATGAGG ATTTTCGAGGA GGAACCCACA TCCCCCATAG GTCAGTGTGT 600
 GGCCATCTAC CACTTTGAAG GGTCCAGCGA GGGCACTATC TCTATGGCCG AGGGTGAAGA 660
 CCTCAGTCTT ATGGAAGAAG ACAAAGGGA CGGCTGGACC CGGGTCAGGC GGAAAGAGGG 720
 AGGCGAGGGC TACGTGCCCC CCTCCTACCT CCGAGTCACG CTCAATTGAA CCC 773

60